On Diagonal Products of Doubly Stochastic Matrices

R. Bapat and T. E. S. Raghavan
Department of Mathematics
University of Illinois at Chicago Circle
Chicago, Illinois 60680

Submitted by Hans Schneider

ABSTRACT

The following result is proved: If A and B are distinct $n\times n$ doubly stochastic matrices, then there exists a permutation σ of $\{1,2,\dots,n\}$ such that $\prod_i a_{i\sigma(i)} > \prod_i b_{i\sigma(i)}$.

A real $n\times n$ matrix $A=(a_{ij})$ is called nonnegative if all the entries a_{ij} are nonnegative. A nonnegative matrix A is called doubly stochastic if its column totals and row totals are unity. We will denote by Ω_n the set of all doubly stochastic $n\times n$ matrices, and by J_n the $n\times n$ matrix all of whose entries are $1/n$. A nonnegative matrix A is said to have doubly stochastic pattern if there exists a doubly stochastic matrix $D=(d_{ij})$ such that $a_{ij}=0$ iff $d_{ij}=0$. Let S_n denote the set of all permutations of $\{1,2,\dots,n\}$. If $\sigma\in S_n$, then the set $\{a_{1\sigma(1)},a_{2\sigma(2)},\dots,a_{n\sigma(n)}\}$ is called a diagonal and $\prod_i a_{i\sigma(i)}$ a diagonal product of A.

We will first prove the following lemma, which slightly generalizes a known result in [6]. Also, it is very closely related to a theorem of Menon [3] and a theorem of Brualdi, Parter, and Schneider (Theorem 6.1) in [1]. The method of proof used here is different.

LEMMA. *Let A,B be nonnegative $m\times n$ matrices such that:*

(i) $A=D_1BD_2$, *where D_1,D_2 are diagonal matrices with positive diagonal entries.*

(ii) $A-B$ *has row totals and column totals zero.*

Then $A=B$.

Proof. Let $D_1=\operatorname{diag}(x_1,x_2,\dots,x_m)$ and $D_2=\operatorname{diag}(y_1,y_2,\dots,y_n)$. Let $C=(c_{ij})=(x_iy_j)$, $i=1,2,\dots,m$; $j=1,2,\dots,n$. Since the x_i's and y_j's are positive, it

0024-3795/80/030071+5$01.75

follows that for any two rows R and R' of C, either $R \leqslant R'$ or $R' \leqslant R$ (coordinatewise). Thus, there is a row which weakly dominates all the other rows. Similarly there is a column which is weakly dominated by all the other columns. Without loss of generality let them be the first row and the first column. We therefore have $x_1 y_1 \leqslant x_1 y_j$ for all j and $x_1 y_1 \geqslant x_i y_1$ for all i. By assumption $a_{ij} = b_{ij} x_i y_j$ for all i, j. In case $x_1 y_1 \leqslant 1$, then $a_{i1} = b_{i1} x_i y_1 \leqslant b_{i1} x_1 y_1 \leqslant b_{i1}$ for all i, and further $\sum_i a_{i1} = \sum_i b_{i1}$. Thus the first columns of A and B coincide. If $x_1 y_1 \geqslant 1$, by a similar argument the first rows of A and B coincide. The conditions of the Lemma still hold when the first row or column is deleted, whichever coincides in A and B. By an elementary induction the Lemma follows. ■

THEOREM. *Let A and B be distinct doubly stochastic matrices. Then there exists a $\sigma \in S_n$ such that $\Pi_i a_{i\sigma(i)} > \Pi_i b_{i\sigma(i)}$.*

Proof. Suppose if possible that

$$\prod_i a_{i\sigma(i)} \leqslant \prod_i b_{i\sigma(i)} \qquad \text{for all} \quad \sigma \in S_n. \tag{1}$$

Since any positive entry of a doubly stochastic matrix lies on a positive diagonal, we have $a_{ij} = 0$ whenever $b_{ij} = 0$. Let

$$c_{ij} = \begin{cases} \log a_{ij} - \log b_{ij} & \text{if} \quad a_{ij} \neq 0, \\ -N & \text{if} \quad a_{ij} = 0, \end{cases}$$

where N is a large positive number chosen to satisfy the following condition: If $\sigma, \theta \in S_n$, $\Pi_i a_{i\sigma(i)} > 0$, $\Pi_i a_{i\theta(i)} = 0$, then $\sum_i c_{i\sigma(i)} > \sum_i c_{i\theta(i)}$. Since S_n has only finitely many elements, this is clearly possible.

As in [4], consider the linear programming problem: maximize $\sum_i \sum_j p_{ij} c_{ij}$ subject to $P = (p_{ij}) \in \Omega_n$, and its dual: minimize $\sum_i u_i + \sum_j v_j$ subject to $u_i + v_j \geqslant c_{ij}$ for all i, j. Both problems are clearly feasible, and hence by the duality theorem they have optimal solutions. Let u_i^*, v_i^*, $i = 1, 2, \ldots, n$, be a solution to the dual problem. Then

$$u_i^* + v_j^* \geqslant c_{ij} \qquad \text{for all } i, j$$

and

$$\sum_i (u_i^* + v_i^*) = \max_{p \in \Omega_n} \sum_i \sum_j p_{ij} c_{ij}.$$

By the Birkhoff–von Neumann theorem,

$$\max_{p\in\Omega_n}\sum_i\sum_j p_{ij}c_{ij}=\max_{\sigma\in S_n}\sum_i c_{i\sigma(i)}.$$

By our choice of N and by (1),

$$\max_{\sigma\in S_n}\sum_i c_{i\sigma(i)}\leqslant 0.$$

Hence

$$\sum_i (u_i^*+v_i^*)\leqslant 0.$$

Let $x_i=e^{u_i^*}$ and $y_i=e^{v_i^*}$, $i=1,2,\dots,n$. Then $x_iy_j\geqslant e^{c_{ij}}$ for all i,j and

$$\prod_i x_iy_i\leqslant 1. \tag{2}$$

Since $e^{c_{ij}}=a_{ij}/b_{ij}$ if $a_{ij}>0$, we have

$$a_{ij}\leqslant x_iy_jb_{ij}\qquad\text{if}\quad a_{ij}>0.$$

If $a_{ij}=0$, then $a_{ij}\leqslant x_iy_jb_{ij}$ trivially holds. Hence

$$a_{ij}\leqslant x_iy_jb_{ij}\qquad\text{for all } i,j. \tag{3}$$

From (3),

$$\sum_j\frac{a_{ij}}{y_j}\leqslant x_i,\qquad i=1,2,\dots,n. \tag{4}$$

By the generalized arithmetic mean–geometric mean inequality,

$$x_i\geqslant\prod_j\left(\frac{1}{y_j}\right)^{a_{ij}},$$

and therefore

$$\prod_i x_i \geqslant \prod_i \prod_j \left(\frac{1}{y_j}\right)^{a_{ij}}$$

$$= \prod_j \left(\frac{1}{y_j}\right)^{\Sigma_i a_{ij}}$$

$$= \prod_j \left(\frac{1}{y_j}\right).$$

Hence

$$\prod_i x_i y_i \geqslant 1. \tag{5}$$

Combining (2) and (5), $\Pi_i x_i y_i = 1$. This implies that equality must hold in (4) for each i, and further, $a_{ij} = x_i y_j b_{ij}$ for all i, j. Therefore $A = D_1 B D_2$, where $D_1 = \mathrm{diag}\{x_1, \ldots, x_n\}$ and $D_2 = \mathrm{diag}\{y_1, \ldots, y_n\}$. Now by Lemma 1 it follows that $A = B$, a contradiction. This completes the proof. ∎

The analogous statement for diagonal sums follows rather easily. Let $A, B \in \Omega_n$ and suppose that $\Sigma_i a_{i\sigma(i)} \leqslant \Sigma_i b_{i\sigma(i)}$ for all $\sigma \in S_n$. Since $\Sigma_\sigma \Sigma_i a_{i\sigma(i)} = \Sigma_\sigma \Sigma_i b_{i\sigma(i)} = n!$, we have $\Sigma_i a_{i\sigma(i)} = \Sigma_i b_{i\sigma(i)}$ for all $\sigma \in S_n$. By direct computation, it is easily seen that $a_{ij} = b_{ij} + x_i + y_j$ for some x_i, y_j. Since A and B are doubly stochastic, $x_i = 0$ and $y_i = 0$ for all i and hence $A = B$.

It is an easy consequence of the theorem that any doubly stochastic matrix A other than J_n has diagonals σ and τ such that $\Pi_i a_{i\sigma(i)} > (1/n)^n$ and $\Pi_i a_{i\tau(i)} < (1/n)^n$. The first half of this statement is contained in [2].

It is to be remarked that we essentially use the fact that A and B are doubly stochastic. At best we can generalize the theorem trivially to the case where all the row totals and column totals in A and B are identically a constant α and A and $B \geqslant 0$. For example, the theorem is not true in general for the case $A, B \geqslant 0$ where $A - B$ has zero row totals and zero column totals. The following is a counterexample:

$$A = \begin{bmatrix} 0.1 & 0.1 & 3.8 \\ 0.1 & 0.1 & 3.8 \\ 3.8 & 3.8 & 0.4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}.$$

Saunders and Schneider [5] use the Gordon-Stiemke theorem to prove several theorems on multiplicative scaling of matrices. Since Stiemke's theorem is closely related to the duality theorem of linear programming, which we use to prove our result, it is perhaps possible that one could find an alternative proof of our result using the theorems in [5].

We sincerely thank Professor Schneider for drawing our attention to several closely related earlier results in this area. We also thank him for suggesting possible improvements in our earlier version of the Lemma.

REFERENCES

1 R. A. Brualdi, S. V. Parter, and H. Schneider, The diagonal equivalence of a non-negative matrix to a stochastic matrix, *J. Math. Anal. Appl.* 16:31–50 (1966).
2 M. Markus and H. Minc, Some results on doubly stochastic matrices, *Proc. Amer. Mat. Soc.* 13:571–579 (1962).
3 M. V. Menon, Matrix links, an extremization problem, and the reduction of a nonnegative matrix to one with prescribed row and column sums, *Canad. J. Math.* 20:225–232 (1968).
4 T. E. S. Raghavan, Some remarks on matrix games and nonnegative matrices, *SIAM J. Appl. Math.* 36:83–85 (1979).
5 B. Saunders and H. Schneider, Cone graphs and optimal scalings of matrices, *Linear and Multilinear Algebra*, to appear.
6 R. Sinkhorn and P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices, *Pacific J. Math.* 21:343–348 (1967).

Received 4 May 1979; revised 20 July 1979